

Fall 1991

## Notes on the Structure of a Theory of Excuses

Michael Corrado

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>



Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

---

### Recommended Citation

Michael Corrado, Notes on the Structure of a Theory of Excuses, 82 J. Crim. L. & Criminology 465 (1991-1992)

This Criminal Law is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

# CRIMINAL LAW

## NOTES ON THE STRUCTURE OF A THEORY OF EXCUSES

MICHAEL CORRADO\*

### I. INTRODUCTION

An enormous amount of literature on legal excuses has developed over the last two decades,<sup>1</sup> most of it devoted to trying to an-

---

\* Associate Professor of Law, University of North Carolina at Chapel Hill.

<sup>1</sup> See, e.g., Michael D. Bayles, *Character, Purpose, and Criminal Responsibility*, 1 LAW & PHIL. 5 (1982); Richard B. Brandt, *The Insanity Defense and the Theory of Motivation*, 7 LAW & PHIL. 123 (1988) [hereinafter Brandt, *The Insanity Defense*]; Richard B. Brandt, *A Motivational Theory of Excuses in the Criminal Law*, in 27 NOMOS, CRIMINAL JUSTICE 165 (J. Roland Pennock & John W. Chapman eds., 1985) [hereinafter Brandt, *A Motivational Theory of Excuses*]; Joshua Dressler, *Justifications and Excuses: A Brief Review of the Concepts and the Literature*, 33 WAYNE L. REV. 1155 (1987) [hereinafter Dressler, *Justifications and Excuses*]; Joshua Dressler, *New Thoughts About the Concept of Justification in the Criminal Law: A Critique of Fletcher's Thinking and Rethinking*, 32 U.C.L.A. L. REV. 61 (1984) [hereinafter Dressler, *New Thoughts*]; Albin Eser, *Justification and Excuse*, 24 AM. J. COMP. L. 621 (1976); GEORGE P. FLETCHER, *RETHINKING CRIMINAL LAW* (1978) [hereinafter FLETCHER, *RETHINKING CRIMINAL LAW*]; George P. Fletcher, *The Individualization of Excusing Conditions*, 47 S. CAL. L. REV. 1269 (1974); George P. Fletcher, *The Right Deed For The Wrong Reason: A Reply to Mr. Robinson*, 23 U.C.L.A. L. REV. 293 (1975); George P. Fletcher, *Should Intolerable Prison Conditions Generate A Justification or an Excuse for Escape?*, 26 U.C.L.A. L. REV. 1355 (1979) [hereinafter Fletcher, *Intolerable Prison Conditions*]; George P. Fletcher, *The Unmet Challenge of Criminal Theory*, 33 WAYNE L. REV. 1439 (1987); Kent Greenawalt, *Conflicts of Law and Morality—Institutions of Amelioration*, 67 VA. L. REV. 177 (1981); Kent Greenawalt, *Distinguishing Justifications from Excuses*, 49 LAW & CONTEMP. PROBS. 89 (1986) [hereinafter Greenawalt, *Distinguishing Justifications from Excuses*]; Kent Greenawalt, *The Perplexing Borders of Justification and Excuse*, 84 COLUM. L. REV. 1897 (1984) [hereinafter Greenawalt, *Perplexing Borders*]; Jerome Hall, *Comment on Justification and Excuse*, 24 AM. J. COMP. L. 638 (1976); H.L.A. Hart, *Legal Responsibility and Excuses*, in DETERMINISM AND FREEDOM IN THE AGE OF MODERN SCIENCE (Sidney Hook ed., 1961); Andrew von Hirsch & Nils Jareborg, *Provocation and Culpability*, in RESPONSIBILITY, CHARACTER, AND EMOTIONS 241 (Ferdinand Schoeman ed., 1987); Sanford H. Kadish, *Complicity, Cause and Blame: A Study in the Interpretation of the Doctrine*, 73 CAL. L. REV. 323 (1985); SANFORD H. KADISH, *Excusing Crime*, 75 CAL. L. REV. 275 (1987); Michael S. Moore, *Causation and the Excuses*, 73 CAL. L. REV. 1091 (1985); Michael S. Moore, *Choice, Character, and Excuse*, 7 SOC. PHIL. & POL'Y 29 (1990) [hereinafter Moore, *Choice, Character, and Excuse*]; MICHAEL MOORE, *LAW AND PSYCHIATRY* (1984); Paul H. Robinson, *Criminal Law Defenses: A Systematic Analysis*, 82 COLUM.

swer two questions: (1) Is there a single theory underlying the excuses that the law recognizes? (2) How are *justifications* related to excuses in the law? The practical import of the first question is clear; if we can find the underlying theory, then not only can we regularize the scheme of existing excuses, we can also decide about *proposed* excuses—about whether, for example, the causal influence of an extra Y chromosome ought to count as an excuse. Or, if that is too ingenuous a view of the role that theory can play, we can at least *consider* each proposed excuse in the light of the most plausible theories. The better view about normative theories appears to be that the process of theorizing proceeds by comparing theory with evidence, modifying our beliefs about one or the other (whichever appears to be the less well-entrenched) until we reach a state of *reflective equilibrium*.<sup>2</sup> In the case of excuses, the evidence would be our considered judgments about what should and what should not count as an excuse in the law. Therefore, a more conservative statement of the significance of the first question would be that the attempt to find a unifying rationale for excuses should help us to clarify our views on newly proposed excuses.

On the other hand, the significance of the second question is far from clear. In the past, lawyers and philosophers saw little reason to distinguish justification and excuse. Recently, however, a great deal of effort has gone into showing that there are practical consequences of distinguishing the two.<sup>3</sup> George Fletcher, for example, argues that where the action one person takes against another is justified rather than excused, that fact has consequences for the rights

---

L. REV. 199 (1982); Paul H. Robinson, *A Theory of Justification: Societal Harm as a Prerequisite for Criminal Liability*, 23 U.C.L.A. L. REV. 266 (1975) [hereinafter Robinson, *A Theory of Justification*]; Richard A. Wasserstrom, *H.L.A. Hart and the Doctrine of Means Rea and Criminal Responsibility*, 35 U.CHI. L. REV. 92 (1967). Glanville Williams, *The Theory of Excuses*, 1982 CRIM. L. REV. 732.

<sup>2</sup> See JOHN RAWLS, *A THEORY ON JUSTICE* 49 (1971):

Moral philosophy is Socratic: we may want to change our present considered judgments once their regulative principles are brought to light. . . . A knowledge of these principles may suggest further reflections that lead us to revise our judgments. This feature is not peculiar though to moral philosophy, or to the study of other philosophical principles such as those of induction and scientific method. For example, while we may not expect a substantial revision of our sense of correct grammar in view of a linguistic theory the principles of which seem especially natural to us, such a change is not inconceivable, and no doubt our sense of grammaticality may be affected to some degree by this knowledge. But there is a contrast, say, with physics. To take an extreme case, if we have an accurate account of the motions of the heavenly bodies that we do not find appealing, we cannot alter these motions to conform to a more attractive theory.

<sup>3</sup> On the side of philosophy, all of this started with J. L. Austin's, *Plea for Excuses*, 57 PROC. OF THE ARISTOTELIAN SOC'Y 1 (1956).

of third parties.<sup>4</sup> For example, Fletcher would say if I am justified in attacking you (in self-defense, say), then a third party could not be justified in stopping me. If I am merely excused for what I do, on the other hand, as when I act under duress from another party, interference by a third party may well be justified.<sup>5</sup> That would be a significant consequence of the distinction if it were true, and it would justify a search for the principles underlying the distinction. But whether third party rights do in fact distinguish the two is a matter of controversy, and it is not clear that there are any other practical consequences of the distinction.

As to the first of these questions, whether there is a single theory underlying the excuses that the law recognizes, it seems to me that the interrelationship of the various theories has never been systematically studied. A careful comparison shows that there are several different questions that these theories have been trying to answer. It is not right to consider them all as conflicting attempts to answer the same question, as writers commonly do.<sup>6</sup> In particular the "voluntariness" theory—we excuse in circumstances in which the actor had no real choice—and the "character" theory—we excuse in circumstances that prevent the inference from the illegal action to the conclusion that the actor has a bad character—attempt to answer different questions. These two theories have been proposed as competing theories of excuse—indeed, as the only plausible competitors. The fact of the matter is that these theories answer different questions: the voluntariness theory sets out a unifying mark of excuses, and the character theory explains why that mark is significant. One general theory might well include both of them; indeed, under uniform assumptions, each of them may require the other. The telling distinction, in any event, is not between a voluntariness theory and the character theory, but between two versions or interpretations of the voluntariness principle, the currently fashionable one and an older, more traditional one.

On the second question, how justifications are related to excuses in the law, a good case for distinguishing justifications and excuses has never been made, and it may be time to abandon the attempt. There is enough to say about justification as a *kind* of excuse and about the features that distinguish it from other sorts of

---

<sup>4</sup> See Fletcher, *Intolerable Prison Conditions*, *supra* note 1, at 1360.

<sup>5</sup> *Id.*

<sup>6</sup> For example, Joshua Dressler, in both his book *UNDERSTANDING CRIMINAL LAW* 183-186 (1987) and in his article, *Justifications and Excuses*, *supra* note 1, lists utilitarian theories, character theories, causal theories, voluntariness theories and others as competing theories of excuse.

excuse. It is one thing to agree that justifications are different from other sorts of excuse, and another to argue that the distinction is so fundamental as to require different underlying theories and to entail different treatment in the law. Academic writers have argued that excuses and justifications are different in the second, more fundamental, way.<sup>7</sup>

In addition to the effect on the rights of third parties mentioned above, academic writers say there are these distinctions to be made: that while excuses have to do with the internal state of the actor, justifications have to do with the consequences of the action; that excused actions should be disapproved though excused while a justified action is cause for celebration; and that the excused actor should regret what she has done while the justified actor should not. All of these propositions appear to me to be false. There is only one reason for either excusing or justifying an illegal act, and that is that the actor does not deserve to be punished. The grounds of desert always depend on the internal state of the actor. This position has apparently been accepted by the authors of the Model Penal Code, and it is a defensible position.<sup>8</sup> Justification is a kind of excuse, and any theory of excuses should be rich enough to account for it.

In this paper I undertake first to compare the theories of excuse that have recently been generated, and then to examine some proposed distinctions between justification and excuse.

## II. CHARACTER, CAUSE, AND CHOICE

I discern at least three, and perhaps four, questions to be answered by a theory of excuses. At the lowest level is the question: Which sets of conditions will excuse illegal behavior? The answer to this question will be entailed by the theory, and it in turn provides a basis on which the theory is to be evaluated. Duress, necessity, insanity, and involuntary intoxication are all generally agreed to be excuses. Any acceptable theory should entail all of these excuses. It is over *proposed* excuses that there is disagreement; whether, for example, a background of deprivation and abuse can ever serve as an excuse.<sup>9</sup> It is because of this that a theory of excuse can have practi-

---

<sup>7</sup> Kent Greenawalt has in fact suggested ways in which justification and excuse are similar but, as I point out below, seems to have stopped just short of arguing that the distinction can be dispensed with. See Greenawalt, *Perplexing Borders*, *supra* note 1; Greenawalt, *Distinguishing Justifications from Excuses*, *supra* note 1.

<sup>8</sup> See *infra* notes 30-31 and accompanying text.

<sup>9</sup> See generally, David L. Bazelon, *The Morality of the Criminal Law*, 49 S. CAL. L. REV. 385 (1976); Stephen J. Morse, *The Twilight of Welfare Criminology: A Reply to Judge Bazelon*, 49 S. CAL. L. REV. 1247 (1976); Richard Delgado, "Rotten Social Background": Should the

cal significance.

At the next level is the question: Is there a feature, or set of features, that all excusing conditions have in common? As I see it, for example, the inability to do otherwise is proposed by some as a general unifying feature shared by all excuses: one who operated under duress, or necessity, or intoxication could not have done otherwise. But while this feature is meant to be something that we find in all excusing conditions, a *criterion* of excuse, so to speak, it does not by itself explain why it should excuse, illegal behavior. What is it about the inability to do otherwise that makes us think we should be more lenient toward those who could not do otherwise?

Thus, a theory of excuse must answer at least one more question: Why should conditions that exhibit the criterial feature excuse? This is a question about the underlying *rationale* for excuses. Utilitarianism is one of the answers to this question; it says that punishment is called for only when it will increase happiness, and where it is pointless to punish behavior ought to be excused. A theory may require another step; it may have to provide an explanation of what ties the rationale to the criterion: why, given the rationale, is this particular criterion significant? Let us call this the question of the *account* to be given of the relation of the criterion to the rationale. Thus, a motive-utilitarianism may pose the mediation of character to explain why certain features excuse. In that case, the theory will have attempted to answer four separate questions about excuses.

A complete theory might hold that punishment is not called for unless it is required to promote an increase in happiness (the rationale).<sup>10</sup> It might hold further that punishment would do nothing to forward that end when the agent who performed an illegal act could not have done otherwise (the criterion). The reasoning is that in those cases, the illegal act is no evidence that the agent has a bad character or a bad set of motives, and thus no evidence that punishment might improve her character by raising her motivation to do things that promote happiness (the account). This is a theory that is utilitarian at the highest level, a character theory at the next level, an "inability to do otherwise" theory at the next. What this theory might entail at the lowest level, about particular excuses, is its basis for evaluation. A theory that entailed that none of the existing excuses—like duress and insanity—should excuse would be rejected,

---

*Criminal Law Recognize a Defense of Severe Environmental Deprivation?*, 3 LAW & INEQ. J. 9 (1985).

<sup>10</sup> See Brandt, *A Motivational Theory of Excuses*, *supra* note 1. See also, Brandt, *The Insanity Defense*, *supra* note 1; Richard B. Brandt, *A Utilitarian Theory of Excuses*, 78 PHIL. REV. 337 (1969).

no matter how appealing at the higher levels. On the other hand, a theory that is successful in accounting for existing excuses will have to be taken seriously in what it says about proposed excuses.

The following list illustrates, in schematic form, the parts a theory of excuse might have:

#### A THEORY OF EXCUSES

- I. RATIONALE: Utilitarian
- II. ACCOUNT: Character
- III. CRITERION: Inability to do otherwise
- IV. EXCUSES: Duress, insanity, etc.

The rationale tells us that the end of the theory of excuses is to promote the general welfare. The account shows us how we are to promote the general welfare by limiting punishment to circumstances in which an inference to bad motivation or bad character is possible. The criterion is a claim about circumstances in which an inference to character is possible. Finally, the excuses are those roughly grouped sets of circumstances which fit the criterion.

#### A. CONFLICT AMONG THEORIES

A character theory—or better, a theory that has such a character element—does not have to be a utilitarian theory. A retributive theory of excuses (perhaps a misnomer for a theory of excuses) might assert that desert is the only appropriate basis for punishment. Only bad character deserves punishment, on such a theory, and we excuse cases in which the agent could not have done otherwise because the inability to do otherwise prevents the inference from the illegal act to the bad character.<sup>11</sup> To think that utilitarianism or retributivism are in competition with character as theories of excuse is just a mistake; and neither is character in competition with the inability to do otherwise.

Naturally some proposals do conflict with others; and once we have lined them up and sorted out the levels at which they are supposed to operate it is easy enough to see which of them conflict with which others. As a rationale, utilitarianism conflicts with retributivism; different sorts of utilitarianism conflict with each other; and different sorts of retributivism conflict with each other. Similarly there are a number of accounts that may be understood as alternatives to the character account: a theory that sees criminal action as an expression of disrespect for the rights of others may account for excuses (whatever the criterion) as cases in which such disrespect

---

<sup>11</sup> See FLETCHER, *RETHINKING CRIMINAL LAW*, *supra* note 1, at 799-802. Fletcher himself notes a fatal defect in this theory, but does not abandon it.

cannot be inferred.<sup>12</sup> Again, the account may be that under conditions set out by the criterion it is impossible to infer that an unfair advantage has been taken.<sup>13</sup> In either case, as in the case of the character account, the underlying rationale may be either retributive or utilitarian.

The "inability to do otherwise" criterion is in competition with what has been called the "causal" theory. A better way to put it may be that there are two interpretations of the inability to do otherwise, yielding two theories that we may call the choice theory and the causal theory. The distinction between the two theories is meta-physical, but it may be the most significant distinction in the whole theory of exculpation. If my argument is sound, our attitudes about a whole range of proposed excuses, and about the relationship of the character theory with the inability to do otherwise, will depend on whether we adopt the causal theory or the choice theory. In this section I will first contrast the causal and choice theories and their consequences, and then show how the choice theory relates to the character theory.

A theory of excuses that involved the "inability to do otherwise" thesis as a criterion of excuses would have to include something like the following proposition:

C1. Someone who has broken the law should be punished if, and only if, she had the capacity and a fair opportunity to have done otherwise.<sup>14</sup>

---

<sup>12</sup> See Benjamin B. Sendor, *Crime as Communication: An Interpretative Theory of the Insanity Defense and the Mental Elements of Crime*, 74 GEO. L.J. 1371, 1398-1401 (1986):

Criminal law demands that a person act with adequate respect for [legally protected] interests by acting in a way that avoids injuring them. . . . A wrongdoer's disrespect expressed by committing a wrongful act evokes our anger and prompts us to blame him, both in everyday life and through the criminal justice system. . . . Only a person who appreciates the circumstances, nature, and consequences of his conduct, and who can control his conduct, expresses respect or disrespect through his conduct.

A character account may also be concerned with disrespect, but it will be in terms of the lack of a long-term motivation to respect the rights of others. As I understand Sendor, a particular instance of disrespect may suffice for punishment, and it is our inability to infer such a particular instance (whatever we may believe about the character of the actor) that explains excuses.

<sup>13</sup> Herbert Morris, *Persons and Punishment*, 52 MONIST 475, 478 (1968):

[P]rovision is made for a variety of defenses, each one of which can be said to have as its object diminishing the chances of forcibly depriving a person of benefits others have if that person has not derived an unfair advantage. A person has not derived an unfair advantage if he could not have restrained himself or if it is unreasonable to expect him to behave otherwise than he did.

<sup>14</sup> See H. L. A. Hart, *Negligence, Mens Rea, and Responsibility*, in PUNISHMENT AND RESPONSIBILITY 136, 152 (1968). If we include mistake and ignorance among the excuses, rather than understanding them as denials of an element of the crime, for example, then it is prima facie doubtful whether the inability to do otherwise, however it is to be inter-

There are, however, two importantly different interpretations of C1 that have been advanced; they arise from two different ways of construing the phrase "The actor could have done otherwise." These two interpretations, which result in what have been called the "choice" theory and the "causal" theory, agree that it is the absence of capacity and fair opportunity to do otherwise that is the feature common to all excuses; but the first presumes a compatibilist understanding of the ability to do otherwise, while the second presumes an incompatibilist understanding of it.

### 1. *A Digression on Compatibilism*

Explaining compatibilism is a complicated matter; but it is now a part of the discussion of legal excuses and cannot be avoided. Readers familiar with this discussion should skip ahead to the next section.

The easiest way of explaining compatibilism is to start with the problem that compatibilism tries to solve, a problem involving determinism. Determinism is the view that every event has a cause. Since a cause is a sufficient condition in the circumstances in which it occurs for the occurrence of the effect, determinism entails that every event is the necessary outcome of the immediately preceding state of affairs.<sup>15</sup> Accordingly, given the initial state of the universe, everything that happens has to happen; there are no alternatives.

But human actions are events like any other, and if there is no alternative to an action that occurs, then there is no free choice. If

---

preted, is a feature shared by all excuses. Can it explain mistake and ignorance? The person who injures someone through a mistake of fact had a choice and knew that she had a choice, but did not know the consequences of the choices. It is possible to stretch the inability to do otherwise to cover that case. The actor *could* have acquired the necessary knowledge to avoid the accident, but requiring her generally to acquire enough knowledge to avoid all mistakes would be so expensive as to be an unreasonable option. There is no fair opportunity to acquire that much information. That is somewhat less than persuasive, I think. This is not a problem, of course, if mistake and ignorance are not to be treated as excuses; but even if they are what it shows is that an inability to do otherwise is the unifying feature only of a large *class* of excuses.

<sup>15</sup> To say that an event has a cause is to say that there was a sufficient condition for its occurrence. Suppose that conditions ABC cause E to occur. Suppose further that, contrary to the hypothesis, E might have failed to occur even though ABC occurred; that is, F might have resulted instead of E. Then clearly "ABC" is not a full statement of the cause of E; something else—call it "D"—must have occurred to produce E instead of F. Given ABCD, E must occur (assuming it to be a caused event); otherwise we will have to look still further for a complete statement of the cause. It is notoriously hard to formulate the causal relation or to say what it is for one event to be the cause of another; but I think that the thesis of determinism is somewhat easier to formulate. It says roughly that every event that occurs in some temporal state of the universe has a sufficient physical condition in a prior state of the universe. The difficulty lies in saying what it is for a state of the universe to be a sufficient physical condition for an event.

there is no free choice, there is (on the "inability to do otherwise" criterion) no criminal responsibility. So either we have to abandon the idea that anyone is criminally responsible, or we must insist that determinism is false. Thus, if you believe that determinism is true, you must believe that responsibility is an illusion; and if you insist that men and women are responsible, you must assume that determinism is false—an assumption that goes far beyond our ability to know. That is a statement of the apparent problem that compatibilism tries to solve.

The conclusion that we must choose between responsibility and determinism depends on the stated assumption—which we may label *incompatibilism*—that if an action is caused by prior events and so is inevitable, it cannot be free. One way to preserve responsibility without ruling out determinism would be to argue that causation does not rule out freedom of the relevant sort. This is the position known as *compatibilism*, and filling it out naturally requires providing the promised interpretation of "free."

An account of compatibilism may proceed like this: an action is behavior that originates in the actor's will and that results from her effort. Such an action will be constrained if outside forces would have kept the actor from doing otherwise. But the actor will be free if there would have been nothing blocking her from doing something else if she had chosen to. The important point is that it is irrelevant to the question of freedom whether prior forces caused her to will what she did. A cause that acts through the actor's will does not destroy the freedom of her action; but a cause that operates outside her will does.

Suppose, for example, that because of experiences she has had or because of her physical makeup, an actor is incapable of avoiding a certain action, the action of shooting someone; the experiences or the makeup *cause* the action. But they act through an internal mechanism where they have become internal causes, and therefore irrelevant, according to the compatibilist view, to the question whether her action is "free". Now change the scenario: Suppose that this same actor, with her internally determined psychological state, was also externally constrained: that someone forced the gun into her hand and pressed her finger on the trigger, so that had she willed to do otherwise she would not have succeeded. Her alternatives were cut off by external constraints, and so she cannot be said to have acted "freely".

Genuine excusing conditions, according to the compatibilist, are more like the second set of conditions than the first. They are external constraints and therefore they are relevant to the question

of responsibility. Whether or not an agent could have done otherwise, according to the compatibilist, depends on whether, if she had tried to do otherwise, she would have been able to do otherwise. It is immaterial whether her mental state was internally determined and she was unable to try to do otherwise; for all we know, all action and all trying is determined by prior events. By this route, determinism is shown to be compatible with responsibility: even if the internal states that lead to our actions are fully determined, those cases in which there are no external constraints are cases of free and therefore responsible action.

A great many philosophers, it seems, hold that compatibilism is true.<sup>16</sup> What is it that makes compatibilism plausible? Why would anyone suppose that responsibility could be affected by external causes but not by internal ones? One argument goes like this: We talk about responsibility and the lack of responsibility without knowing whether determinism is true. The features that we base responsibility on would still be there if we learned that determinism is true, and they would be there if we learned that determinism is false. Therefore determinism must be irrelevant to questions of responsibility.

I am not sympathetic with this argument; it depends upon assumptions that are too far removed from reality. It asks what would happen if we learned all at once that determinism is true. If we were to make that discovery, it is true, we would not be likely to abandon our existing world-view. Such a psychological inability, though, would have nothing to do with the rightness or wrongness of that world view in light of the new discovery.

It seems more reasonable to ask whether there are any conditions under which we would change our minds about freedom and responsibility. Would a gradual discovery that determinism was likely to be true have an effect on our views? We have learned a fair amount over the last while about the causes of behavior; and as we learn that a piece of behavior is caused by some prior condition, we are inclined to remove that piece of behavior from the class of responsible behavior. What we might have called responsible behavior before learning about brain tumors, for example, we might now classify as caused and as excused behavior. If we ever do learn that general determinism is true, it will be piecemeal; we will learn it by putting together information about different sorts of causes and dif-

---

<sup>16</sup> See DANIEL DENNETT, *ELBOW ROOM: THE VARIETIES OF FREE WILL WORTH WANTING* (1984); R. E. Hobart, *Free Will as Involving Determination and Inconceivable Without It*, 43 MIND 1, *passim* (1934); Moritz Schlick, *When is a Man Responsible?*, in *FREE WILL AND DETERMINISM passim* (Bernard Berofsky ed. 1966).

ferent sorts of effects. If determinism is true, then as our knowledge gradually approaches that understanding, more and more types of behavior are likely to turn out not to be responsible behavior. Naturally, if somehow we learned overnight that it is true we would still be likely (for a while, anyway) to hold people responsible. But our inability to deal with such overwhelming information is no evidence as to what the *correct* response should be.

It is important to point out that neither compatibilism nor incompatibilism entails either determinism or indeterminism. If compatibilism is true, then whether or not determinism is true we are still responsible for most of what we do. If incompatibilism is true, then we are responsible only if determinism is false. Determinism, it seems to me, is for us an empirical question; although philosophers have attempted proofs of its necessary truth or falsity,<sup>17</sup> it is something to be established or refuted (if at all) with the tools of science. The genuinely philosophical question is about whether compatibilism is true or false.

## 2. *Cause and Choice Again*

To return to the "inability to do otherwise" criterion of excuses: The popular version is the compatibilist version, the version I have called the choice theory, which takes lack of freedom—the inability to do otherwise—as the mark of excuse, but counts only external constraints as limiting freedom. The other, incompatibilist version, is what I have called the causal theory. I call it the causal theory because it recognizes the effect of causation on responsibility, whatever route that causation takes, whether internal or external. It entails, among other things, that if determinism is true, then everything should be excused.

These two theories entail different conclusions about what excuses ought to be admitted. Both account (by and large) for the existing excuses,<sup>18</sup> but they disagree about the newly proposed excuses. For example, recall that social background has been proposed as an excuse: if a causal chain can be established between a deprived and abusive background and a particular crime in a particular case, then the crime ought to be excused—since causation of an action means the actor could not have done otherwise. But that argument clearly presupposes incompatibilism and the causal theory. On the choice version, the mere fact that the action was caused

---

<sup>17</sup> See PETER VAN INWAGEN, *AN ESSAY IN FREE WILL* (1984).

<sup>18</sup> But see Michael Corrado, *Automatism and the Theory of Action*, 39 EMORY L.J. 1191 (1990) (arguing that compatibilism gives the wrong answer in connection with automatism and related defenses).

doesn't mean that the agent could not have done otherwise. If she was internally rather than externally constrained, then on the compatibilist view of the choice theory she can be said to have acted freely and to be responsible for what she has done. It therefore follows from the causal theory but not from the choice theory that causation by background may sometimes be an excuse.

These two interpretations of the "inability to do otherwise" criterion of excuse have radically different implications for the role of character in a theory of excuses.

#### B. THE EQUIVALENCE OF CHOICE AND CHARACTER

A theory of excuses that has a character account of the relation between the rationale and the criterion will include a proposition like the following:

- A1. Someone who has broken the law should be punished only if the circumstances permit an inference from the action to a defect in the actor's character.

Thus, if the circumstances do not permit an inference to a defect in character, the actor must be excused. What is a character defect? According to one writer, it is an inadequate motivation to do the right thing.<sup>19</sup> An excuse is a state of affairs that prevents us from inferring that an action is evidence of such inadequacy. Someone who, at the point of a gun, throws a rock through a neighbor's window may nevertheless have an adequate amount of motivation to respect her neighbor's property. *Adequate* motivation is not the same as *blind* or *absolute* motivation; we do not expect someone to sacrifice her own life to save her neighbor's windows. Therefore, the fact that the sacrifice of her life may have appeared to her, the only other alternative prevents us from concluding that she does not have the right amount of motivation, or that her character is defective.

I have already given the general formulation for an "inability to do otherwise" criterion of excuse: Someone who has unjustifiably broken the law should be punished if, and only if, she had the capacity and a fair opportunity to have done otherwise. In the example above, a gun at the actor's head deprived her of a fair opportunity to do otherwise. Intoxication, on the other hand, might deprive someone of the capacity to make a choice. According to this theory, the action will be excused if either the opportunity or the capacity is absent.

These two theories, the character account and the inability-to-do-otherswise criterion, do not compete with one another but (as I

---

<sup>19</sup> See Brandt, *A Motivational Theory of Excuses*, *supra* note 1, at 165, 174.

have already argued) are answers to different questions. I have illustrated how a general theory might actually combine the "inability to do otherwise" criterion with a character account. But keeping in mind the two different interpretations of the "inability to do otherwise," it is possible to reach a more surprising conclusion: The incompatibilist causal version of the inability to do otherwise actually *does* exclude a character account. It is only the compatibilist "choice" version that is consistent with the character account. In addition, a general theory that contains the compatibilist "choice" version of the inability to do otherwise as a criterion of excuses may be *equivalent* to one that contains the character theory as an account of the relationship between the criterion and the rationale.

First, we may show that the causal theory does exclude a character account. The causal theory, remember, is an interpretation of the "inability to do otherwise" theory that requires indeterminism before it will say that an action has been free and responsible.<sup>20</sup> This theory appears to be genuinely inconsistent with the character theory. Character, on the causal theory, is something over which we have only the most minimal control since control requires uncaused choice.<sup>21</sup> If our bad actions are caused by something over which we have no control, then we are not responsible for them. In the character theory, on the other hand, it is precisely those things that are caused by our characters without the interference of outside causes that we are responsible for. A general theory that includes the character account, therefore, would seem incompatible with a general theory that includes the causal criterion.

To show that choice and character are equivalent, on the other hand, requires showing that each entails what the other entails. Though it would be impossible here to do that formally (we do not even have a formal characterization of what a general theory of excuses would be), it is easy enough to do in a rough and practical way. A theory that includes a character account will count as an excuse anything that blocks an inference from action to character.

---

<sup>20</sup> "Undetermined" and "indeterminate" do not mean random; in fact someone whose actions are undetermined may nevertheless be influenced by her circumstances in making her choice. That is, after all, what we appear to mean by rational choice, and no one has shown that undetermined but influenced action is a logical impossibility.

<sup>21</sup> The most common objection to incompatibilism is that uncaused choice is random choice—hardly a basis for ascriptions of responsibility. That argument seems to me to beg the question; if "random" is to have any bite here, it must mean something like "not ascribable to the actor." But that uncaused action cannot be ascribed to the actor is a conclusion, not a premise, and I have not yet seen the argument for that conclusion. Incompatibilism nevertheless has problems of its own; the long literature shows how difficult it is not only for the compatibilists but also for the incompatibilists to explain what it means to say that someone could have done otherwise.

Thus any circumstance that makes it impossible or sufficiently difficult for the actor to have done otherwise will be an excuse, because in those circumstances the action will not reveal her character. There is, however, one qualification: the fact that a piece of behavior was *dictated by* the actor's character may have made it hard or impossible for her to have avoided it, but that will not prevent an inference to her character. Just the reverse is true: it facilitates such an inference, and thus will not be an excuse.

But we should expect very much the same result from the choice criterion. Those actions which follow from internal causes—and the limits of what counts as an internal cause can be made to vary with the limits of what is attributable to character—are not, on the choice theory, actions that the actor was unable to avoid in the absence of outside constraints. Therefore, actions that follow from internal causes are not excused. And thus a character account will entail that some functional equivalent of the inability to do otherwise, in its compatibilist interpretations, will serve as a criterion—at least a sufficient condition—of excuses.

Similar reasoning works in the other direction, to show that the choice criterion will entail some functional equivalent of a character account. The behavior an actor cannot avoid performing is behavior subject to outside constraints (according to the choice criterion); and such constraints do impair the inference to character. But action due to internal causes is not to be excused, according to the same criterion. Indeed, the fact that an action was caused by an internal cause should facilitate an inference to character.

Of course “internally caused” and “in character” do not mean exactly the same thing. It is logically possible that something could be internally caused without being in character.<sup>22</sup> In fact, the two

---

<sup>22</sup> What if, under one of the science fiction hypotheses that philosophers seem so fond of, someone inserted into my brain something that caused me to kill someone? That would be an internal cause, would it not? And yet the action it produced could not be said to have been caused by my character. Doesn't that demonstrate a difference between the two? In point of fact such a circumstance is so clearly an excusing condition that the choice theorist no less than the character theorist will have to argue to show that her theory would count it as an excuse. That means that the choice theorist must deny that the action is internally caused. The insert is an external cause, and thus even on the choice theory interferes with my free choice.

To illustrate this point, consider the following passages from Michael Moore's *Choice, Character, and Excuse*. In discussing the implications of his own choice theory, which he presents as an alternative to the character theory, Moore tries to show why actions performed under the influence of certain emotions are not for that reason excused:

If all of this is true, then it may seem that . . . internal factors like emotions cannot be said to incapacitate our choices, except by an impermissibly narrow view either

phrases are used with so little precision in the debate over excuses that almost any relationship between them may be argued to be logically possible. Still, they are clearly intended to cover the same ground. Whatever is caused by character will be something the compatibilist will want to treat as internally caused. And to be internally caused is to be caused by the actor's character. To say that an action is out of character when it occurs without outside constraints is, on the view of character theorists, like saying of a piece of wood that catching fire when heat is applied to it is no part of its character. Only those actions that happen under external constraint are out of character, and thus excused.

The upshot of all this is that a choice theorist, whether she does so wittingly or not, will carve out a functional equivalent of character, consisting of all those things that the compatibilist calls *internal* causes. A character theory will have the functional equivalent of a choice criterion in that it focuses on those circumstances that prevent an inference to character. Where the character theorist says, "[I]t is unfair to hold people responsible for actions which are out of character, [but] fair to hold them so for actions in which their settled dispositions are centrally expressed,"<sup>23</sup> the choice theorist should, *mutatis mutandis*, say, "It is unfair to hold people responsible for actions caused or constrained by forces external to them, but fair to

---

of who we are or of what our choosing agency consists [i.e., by impermissibly treating the emotions as external causes].

Moore, *Choice, Character and Excuse*, *supra* note 1, at 39. That is, whether emotions interfere with choice (on the compatibilist view) depends on whether or not they are internal, i.e., part of *us*. To treat them as incapacitating is to treat them as external. Since they are not properly excuses, they should be viewed as part of the person making the choice. Later, discussing the character theory, he says this:

As we saw earlier, the choice theorist had to decide whether an emotion . . . was part of [the actor's] self, and if it was, whether such emotion was part of that part of [the actor's] self that I called [the actor's] choosing agency. The character theorist faces a similar decision.

*Id.* at 43. He admits that one "commonly perceived" difference between character and choice does not exist—that strong emotion should excuse on the choice view and not on the character view:

[T]his difference can only exist if we make asymmetrical assumptions about the two theories. . . . [I]f we allow each theorist the most plausible assumptions about what choosing [interpreted in a compatibilist way] and character include, this difference disappears.

*Id.* at 44 (emphasis added). Moore argues nevertheless that the two theories are distinct. It seems to me, however, that each theory is pliable enough to be fitted within the same constraints and molded to the excuses we want to generate. It also seems that when we do put the same requirements on them—that is, place them under symmetrical assumptions—and mould them to fit a given set of excuses, far from being alternative theories that exclude one another, a general theory containing one of them will also contain the other.

<sup>23</sup> NICOLE LACEY, STATE PUNISHMENT 68 (1988), quoted in Moore, *Choice, Character, and Excuse*, *supra* note 1, at 51.

hold them so for actions caused or constrained by internal forces." The two (depending on the interpretation of "internal" and "character") may well come to the same thing.

The best way to show that this conclusion is true is to consider an argument that has been offered against it. Michael Moore, in attempting to show that the two theories are incompatible, argues that they reach different conclusions about these two cases:

- (1) An actor freely chooses to do wrong (i.e., he had capacity and a fair opportunity not so to choose), and yet the action is 'out of character;' and
- (2) an actor's behavior is good evidence for a settled disposition, and yet he has not (yet) chosen to act on that disposition.<sup>24</sup>

The character theorist, he asserts, will give the wrong answer to each (punish the second but not the first), while the choice theorist will give the right answer to each (punish the first but not the second). If Moore is right about this, then I am wrong and the two theories, choice and character, are clearly distinguishable.

Taking Richard Brandt as a representative character theorist (and I know of no one who has given more thought to this matter than Brandt has), an action demonstrates something about character if it demonstrates something about long-term motivation.<sup>25</sup> The crucial question is whether behavior demonstrates insufficient motivation to do the right thing. Although long term motivation, like character, must be evaluated over a period of time, doing the wrong thing once in the absence of outside constraints may indicate a lack of motivation.

Take the person who, having led an exemplary life up till now, sets fire to her neighbor's house freely and under the influence of no external forces. If I understand Moore correctly, he believes that the choice theorist would refuse to excuse because the action was freely chosen, but that the character theorist *should* excuse, if the action is genuinely out of character. But is that in fact the right conclusion? What can we say about someone who freely burns down her neighbor's house? Does she have adequate motivation to respect her neighbor's property? It is easy enough to *say* that one may act freely and yet out of character, but the free choice of a wrong clearly indicates something about the level of an actor's motivation. If more long-term motivation to respect her neighbor's property would have prevented her from doing what she did—that is, if having a different character would have prevented it—then the character theorist must say that punishment is appropriate in this case. It

<sup>24</sup> Moore, *Choice, Character, and Excuse*, *supra* note 1, at 51.

<sup>25</sup> See generally Brandt, *A Motivational Theory of Excuses*, *supra* note 1, *passim*.

is hard to see how an increase in motivation might not have prevented it. I think, therefore, that the character theorist will never be required to admit that some wrong could be freely chosen and yet fail to indicate something about long-term motivation, as Moore's first example requires.

The second example, in which an actor's behavior is good evidence for a settled disposition and still she has not yet chosen to act on it, points out the most serious defect in the character theory. If it is bad character that calls down punishment—either on a retributive view because it deserves it, or on a utilitarian view in order to modify character in the future—then a bad character ought to be punished even if the actor has not yet done anything wrong. Wrong action is only evidence of character, on a character theory, and is not the theoretical basis for punishment. But since no one is entirely responsible for her character, no one ought to be punished for her character. Thus a character theory cannot be correct. Moore's example is intended to distinguish the choice theory from the character theory by showing that the former does not punish people for things they are not responsible for.

The character theorist can avoid this undesirable outcome in practice by insisting that the *only* good evidence of bad character is bad action freely chosen. The distinction Moore is after is a theoretical one, though, about the basis of punishment. The character theorist bases punishment on something over which the actor has no control—her character. Nevertheless, if the practical outcome is the same, if both theories would only punish freely chosen wrong behavior, then Moore's distinction is insignificant. Indeed, the compatibilist seems to be driven to the conclusion that character is freely chosen. If character is determined by the sum of our prior behavior and our reactions to prior experiences, character would seem to be as freely chosen as that previous behavior and those previous reactions were. At least it must be freely chosen if the present behavior that results from those earlier causes can be said to be free, as the compatibilist would insist.

The compatibilist believes that behavior may be free even if the effort of will that leads to it is the inevitable outcome of prior internal causes. In the end, it is the effort of will, and not the resulting behavior, that the choice theorist must punish. The actor has no more or less control over efforts of will than over her character. As the compatibilist defines "control," the actor *may* be said to control her effort of will, but in that sense she also has control over her character.

Moore is right that the *principle* behind the character account is

different from the principle behind the choice account. One relates excuses to character and the other to an effort of will. But that is because they serve different functions in a theory of excuses. Why should punishment depend upon the freedom of the effort of will? Because only when the will has effect without interference will it tell us something about character. The two theories are therefore consistent with one another, and if my argument is correct, they may, in their most plausible versions, actually require one another.

The important distinction between theories of excuse, therefore, is between compatibilist and incompatibilist theories, between choice and causal theories. As I have already indicated, these two sorts of theories will give radically different answers to questions about proposed excuses. The compatibilist has, for example, no time for Judge Bazelon's plea for an excuse based on a deprived and abusive social background. The product of such a background is the very sort of person we want to deter or put away.<sup>26</sup> The incompatibilist, on the other hand, would need only evidence that the criminal behavior was determined by the actor's background to believe that the behavior ought to be excused.

I draw the following conclusions, therefore, about the structure of a theory of excuses. There are at least three and perhaps four questions to be answered by such a theory. Each of the theories that has been advanced so far is really an answer to only one of these questions. Accordingly, not all these theories are incompatible with one another; in particular the choice theory is not incompatible with the character theory, and neither of these is incompatible with utilitarianism. Finally, the choice theory may well be equivalent to the character theory. The choice theory's real competitor is the causal theory. Since my topic is the *structure* of a theory of excuses, it is not my purpose to argue for one of these two. Most writers feel that compatibilism is so obviously true that no argument is needed in its support. They may be right; but it seems to me still a little early to write off the causal alternative.

### III. JUSTIFICATION AND EXCUSE

In the early part of the century it had become commonplace that there is no significant difference between justification and excuse. That changed with J.L. Austin's "Plea for Excuses."<sup>27</sup> It is now taken to be a sign of learning in this area to be able to distinguish the two, and writers and teachers are careful not to use one

---

<sup>26</sup> See Bazelon, *supra* note 9.

<sup>27</sup> See Austin, *supra* note 3.

label where the other is called for. But all of this fuss, and the industry based on trying to say how precisely the two are to be distinguished, seems to me to be rooted in an error. The older view is right; though we *may* call one sort of thing justification and another sort of thing excuse, in point of theory we are only talking about two sorts of excuse. There is nothing wrong with saying, for example, "We excused her because her behavior was justified." It is hardly the egregious error it has been made out to be.

That doesn't mean that there is nothing of theoretical interest to say about justification. There are all sorts of distinctions to be made within the field of excuses, most of them so far hardly noticed. The distinction between justification and other sorts of excuse can most fruitfully be seen as one of these distinctions. In this section I will try first to show why seeing justification as a type of excuse is the better view and what sorts of errors lead to the view that justifications are fundamentally different from excuses. I will then try to show how justification fits into the scheme of excuses.

#### A. JUSTIFICATION AS EXCUSE

Although there is general agreement that justifications and excuses are different sorts of things, there is a surprising amount of disagreement about what are and what ought to be the characteristic features that distinguish justification from excuse. Though there is a loose agreement that justified actions are right or desirable actions, while excused actions are wrong actions for which the actor ought not to be held accountable, there is no agreement about what "right" and "wrong" mean in this context. Is a right action an action that is morally required, the best available alternative; or is it merely one that is morally permissible? Is it necessary that the action actually *be* right, or does it suffice if the agent reasonably believes it to be right? Reflection on these questions will undermine the view that justification and excuse are different things.<sup>28</sup>

##### 1. *The Right Consequences*

It is often said that a justified action is one whose consequences are desirable, or whose consequences are right in the circumstances. Glanville Williams, for example, has said: "Normally a justification is any defense affirming that the act [is], on balance, to be socially

---

<sup>28</sup> Kent Greenawalt has used arguments similar to those that follow to demonstrate the difficulty of distinguishing justifications from excuses. See Greenawalt, *Perplexing Borders*, *supra* note 1.

approved, or [is a matter] about which society is neutral.”<sup>29</sup> In context, actions which conform to the law might have bad consequences. For example, we might have to allow a neighbor’s house to burn if we want to avoid trespassing on her property. In such a case, we are justified in not conforming to the law; but only if what we do is the right action in the circumstances. The problem is in interpreting the word “right.” Must a justified action be morally required, or is it enough that the action be morally permissible?

This question lends itself naturally to an act-utilitarian interpretation: must a justified action be one that creates more good (happiness, wealth) than any alternative, or than any reasonable alternative; or does it suffice that it creates at least as much good as some reasonable alternative? The important point is that it is the objective moral value of the action that is at issue, and not the state of mind of the actor. But other sorts of ethical theory besides utility also admit of comparisons of the external features of an action. One’s obligations, however derived, do not depend on an actor’s state of mind, though whether one ought to fulfill one’s obligations may. If I fail to keep a promise because I have forgotten it, my omission may be excused. If I fail to keep it because of a more important obligation, my omission may be justified. In either case, however, there is an unfulfilled obligation.<sup>30</sup> Whether an action fulfills an obligation is an external feature of the act that does not depend on the actor’s state of mind. Similarly, those who are willing to talk about rights may say that whether an act invades someone’s rights is an external feature of the act and does not depend on the intention or belief of the actor. Nothing in what follows is meant to restrict the discussion to utilitarianism.

For most of the actions we perform, there are alternative actions that we might have performed instead. These alternatives may be graded morally, and may also be graded according to difficulty. Suppose that conforming to the law will produce a bad result. There will be a range of alternative actions with more desirable consequences. Of these, some are reasonable in difficulty, and some are unreasonably difficult; of those whose difficulty is not unreasonable, some are morally acceptable and others are morally unacceptable. Not every alternative to a legal action with bad consequences will be of a reasonable level of difficulty. More importantly, not every alternative that is more desirable (that is, not every lesser evil) will be morally acceptable.

<sup>29</sup> Williams, *supra* note 1, at 735.

<sup>30</sup> Lawrence L. Heintz, *The Logic of Defenses*, 18 AM. PHIL. Q. 243, 243 (1981).

Which, among the alternatives to legal actions with bad consequences, are justified actions? The position called "perfectionism" can be ruled out; to be justified an action need not be the best possible action in the circumstances, since the best possible action may be unreasonably difficult. On the other hand, 'latitudinarianism' will also be false: not every lesser evil that is not unreasonably difficult is justified. If I see someone about to shoot someone else in the head, I may not necessarily hack off the arm of the attacker to prevent the crime even though that would be a more desirable result than the threatened loss of life. If I could accomplish the same end by getting the attention of the policeman on the corner, hacking off the arm would be unacceptable.

Justification lies between these two extremes. Since perfectionism runs afoul of unreasonable alternatives, one possibility is that a justified action must be the best of all the *reasonable* alternatives (that is, of all the alternatives that are not unreasonably difficult). But even that is too high a standard. Think of the case of *Ploof v. Putnam*.<sup>31</sup> Ploof attempted to tie his boat up to Putnam's dock in a storm, and Putnam's servant cast him adrift again. Ploof's property was damaged. Tying up to Putnam's dock without his permission was illegal; yet the court held that Ploof was justified, and that it was wrong for Putnam to hold him off. Ploof was justified by the great harm he was liable to suffer (and in the end did suffer) if he fastidiously obeyed the law.

Now what if Putnam had established that at the time Ploof became aware of the storm he was several feet closer to someone else's dock? Or that he was the same distance from someone else's dock but that that second dock was sturdier and less likely to be damaged by Ploof's boat in the storm? In either of those cases tying up to the second dock would have produced better consequences than tying up to Putnam's dock, and in each case going to the second dock would have been at least as reasonable as going to Putnam's dock. (Remember, here we are concerned with consequences or external features of our actions, and not with our perceptions of those consequences.) Yet it hardly seems likely that the court would have admitted the existence of those admittedly preferable alternatives as evidence that Ploof's action was not justified. It is implausible, for example, that the court would have found it relevant that Putnam's lawyer could present evidence in court that several other docks were a few feet closer to Ploof at the moment he discovered his peril than was his client's.

---

<sup>31</sup> 72 A. 188 (1908).

It will rarely be the case that there is just one action that the actor would have been justified in performing. But if an actor has several acceptable alternatives before her, then none of them is morally required. Each is at most morally permissible.<sup>32</sup> If an action is morally required, then it is morally forbidden not to do it. It follows that where several actions are acceptable, none of them can be required. Where there is a range of morally permissible actions, what is required is that some *one* of those actions be performed. What is forbidden is that an alternative outside that range be performed.

Envision the choices facing Ploof. He could have tried to weather the storm; he could have docked at Putnam's dock (which was strictly illegal); or he could have docked at any of several other docks (equally illegally). Some of the other docks may have been more suitable but were unreasonably far away—he might not have been able to reach them in time. There may have been several like Putnam's, which would have been damaged if Ploof tied up to them but whose likely damage was much less than the threatened damage to Ploof's boat if he didn't tie up at all. Tying up at any of these would thus have been a lesser evil. Finally, there may have been some reasonably accessible but delicate docks which were constructed hurriedly to permit the unloading of emergency medical supplies and which would have withstood the storm but not the added strain of Ploof's boat. The likely damage to any of these latter docks would have resulted in costs less than the likely cost of the destruction of Ploof's boat, but let us also assume that the cost would have been significantly more than the cost of the likely damage to Putnam's dock or any similar dock.

Ploof *cannot* be required to have tried to reach the farthestmost, least damageable docks, because of the great difficulty involved. The acceptable choices in this setting would have been (1) to try to weather the storm; and (2) to dock at any of the reachable docks except those constructed to receive medical supplies. Weathering the storm would have been the greatest evil, considered in isolation; it threatened both life and property. It is acceptable only in the sense that it accords with the law. The other listed alternatives are

---

<sup>32</sup> See Dressler, *New Thoughts*, *supra* note 1, at 82-83. Dressler makes the point that there may be several actions, none better than the other, which are nevertheless all better than any alternative action. Any of them would be justified, but none could be called the right action. Cases in which we are faced with several alternatives that are indistinguishable in this way seldom arise. My point is that even though it may be possible to rank order all the alternatives, there may still be more than one alternative that is justified.

“lesser evils.” Yet not all are morally permissible. “Permissible” and “impermissible” denote the two relevant classes of actions, and nothing else is relevant. In particular, a justified action need not be the best of anything.

Now here is the question: if it is enough that a justified action be one that is permissible, is it possible to distinguish justified actions from excused actions? Does it make sense to say that a justified action is one that is right and socially desirable, while an excused action is one that is wrong and socially undesirable? Are not actions that are excused—actions performed under duress, for example—also considered permissible or acceptable in the circumstances? We might say about such actions that it would be expecting more than a normal person could bear to have required the defendant to obey the law in those circumstances. Does not that mean that we find the behavior, though unfortunate, permissible? We will return to this question below.

## 2. *State of Mind*

Whatever the external feature that is relevant to justification (whether permissibility or preferability or requirement), is it necessary that the action actually have it, or only that the actor reasonably believe the action to have it? For example, suppose that my neighbor’s favorite tree in her yard is diseased, and it presents an imminent danger of toppling over onto a third neighbor’s house causing serious damage. The only way to avoid this outcome is to cut down the tree at once. There is not even time to call a specialist. It has to be done by someone in the area in the next few minutes. The owner, however, cannot be reached. Since the tree is rotten, it has no economic value. Although cutting down a neighbor’s tree without her permission is illegal, under these circumstances it would be a desirable thing to do—even the *right* thing to do in the circumstances, from an external point of view.

I am unaware of all this, though; I am angry with my neighbor and I cut down her favorite tree while she is away simply to spite her. I am later apprehended and charged with the destruction of the tree; but having learned of the facts in the meanwhile, I plead justification. Had I not done what I did, the tree would certainly have caused serious damage far outweighing the economic value of a diseased tree. The external features of the action fit the requirements for the justification of necessity:

Conduct which would otherwise constitute an offense is justifiable and not criminal when . . . [s]uch conduct is necessary as an emergency measure to avoid an imminent public or private injury which is about

to occur by reason of a situation . . . which is of such gravity that, according to ordinary standards of intelligence and morality, the desirability and urgency of avoiding such injury clearly outweigh the desirability of avoiding the injury sought to be prevented by the statute defining the offense in issue.<sup>33</sup>

Should I be acquitted on those grounds? If we attend only to the external features of the act, as the statute seems to require, then it may well be considered desirable. But if we derive justification from that the result is odd. I will be exonerated for destroying someone else's property with a base motive. If we look to motive then the action should not be approved or justified. But then the distinction between justification and excuse begins to fade, for excuse is said to be the defense that looks to facts about the actor's internal state.

There are three standard positions on this question: (1) The act is justified if, and only if, the act has the appropriate external features. This is the pure externalist position. If this position is correct, my tree-cutting escapade is justified. (2) The act is justified if the act has the appropriate external features and I believe that it has those features. This is the modified externalist position. My cutting down the tree would not be justified on that account, but had I known of the disease and the harm it would have caused my act would have been justified.<sup>34</sup> (3) The act is justified if, and only if, the act would have the appropriate external features *if the circumstances were as the actor believed them to be*. This is the internalist position. This would justify my action where I cut down the tree mistakenly believing it to be diseased and an imminent danger to the neighbor's house, where actually neither was the case. The standard setting in which this claim arises is in what is called "putative self-defense," where Jones, reasonably but mistakenly believing that Smith is attacking her, injures or kills Smith in what she believes is self defense.<sup>35</sup> If this third position is correct, then Jones is justified. One version of this position would require mere belief ("pure" internalism), another would require reasonable belief ("modified" internalism), about the circumstances.

On the second of these positions, which is Fletcher's, no action is justified unless it has the right external features (cutting down the tree really did prevent harm) and *also* is accompanied by the right state of mind (I intended to prevent the harm, knew that I would

---

<sup>33</sup> N.Y. PENAL LAW sec. 35.05(2) (Consol. 1987).

<sup>34</sup> There is room here for the standard controversy over whether awareness suffices or intention to do the right thing is required.

<sup>35</sup> Greenawalt, *Perplexing Borders*, *supra* note 1, at 1922, argues that putative self-defense should be a justification, not an excuse. See also Dressler, *New Thoughts*, *supra* note 1, at 81-86.

prevent the harm). But if the first condition is a necessary condition, why is it not also sufficient (as Paul Robinson, for example, believes)? Why not make the external rightness of the act the test for justification? The answer, I believe, is that that position is simply too bizarre to win acceptance. Ploof would be justified in tying up before a storm to Putnam's dock, much to Putnam's later damage, even if he did not know a storm was coming and did it out of malice. Ploof could argue in court: "Luckily for me a storm came up immediately that justified my presence there." There would be no need even to pretend that he was aware a storm was coming. But if Ploof had good reason to believe that there would be a storm and tied up to avoid it, but no storm came, his action would not, according to this view, be justified.

I believe that the we do not make the external features a sufficient condition because there is a strong sense that it would be wrong to exonerate someone who did not have the proper state of mind. It would be wrong to exonerate someone, for example, who killed another out of malice, thereby unwittingly preventing that person from killing others in the course of a rampage. Even though worse harm was prevented, why should the person who killed the madman, perhaps after careful deliberation and with the intention to inherit his estate, and unaware of the danger that she had prevented from being realized, be set free? The injustice of this consequence is simply so strong that it makes Robinson's externalist position impossible to accept. But if the right state of mind is necessary for justification, why is it not also sufficient?<sup>36</sup> Isn't the important thing that each of us act on the best information she has, information that will sometimes be mistaken?

It is apparently the position of the Model Penal Code that it is in fact sufficient. MPC section 3.02, "Justification generally: choice of evils," reads as follows:

- (1) Conduct which the actor believes to be necessary to avoid a harm or evil to himself or to another is justifiable, provided that:
  - (a) the harm or evil sought to be avoided by such conduct is greater than that sought to be prevented by the law defining the offense; . . . .
  - . . . . .

---

<sup>36</sup> This is an area that lends itself to make-weight arguments. Here are two, one on each side of this issue: On the one hand, rewarding by exonerating the externally right act will encourage people to strive for it, but rewarding the right intention alone will encourage dissimulation. On the other hand, rewarding reasonable attempts whether or not successful will encourage reasonable attempts, but rewarding only successful ones will drive people away from situations where difficult duties might arise.

(2) [where the actor negligently brought about the necessity, or negligently appraised the necessity of her behavior, no defense.]

Thus, although her belief must be reasonable, where the actor does have a reasonable but mistaken belief in the necessity of her action, the MPC would justify her behavior.

Although this internalist position is one, as Fletcher says, that ought to appeal to many,<sup>37</sup> not many have come to its defense. The fear may be that if justification depends not at all on what external features an act has but entirely on the actor's (or a reasonable person's) state of mind, then justification will blend imperceptibly into excuse. Excuse includes not only incapacity defenses like insanity, infancy, and involuntary intoxication. After all, excuse includes what may be called the deliberative defenses like duress and provocation in which the actor is faced with genuine choices, one of which (obeying the law) is generally agreed to be too much to ask of someone in her circumstances. But from choices the actor cannot reasonably be called on to make, it is but a step to choices the actor *should not* make, like the choice to respect property when someone's life is at stake. By these steps, on the internalist view, justification becomes simply a variation of excuse, depending entirely on facts about the actor and not at all on the consequences of her action. Since most of those writing in this area are concerned to emphasize the distinction and not to diminish it,<sup>38</sup> it is little wonder that they have not rushed to the internalist position.

Nevertheless, Kent Greenawalt appears to have moved to the internalist position in two recent papers on this topic.<sup>39</sup> In these papers, he tried to show the difficulty of distinguishing justification from excuse, and thus had no stake in affirming the external element in justification. While he has not explicitly adopted the internalist position, he maintains both of these propositions, which together seem to entail internalism: (1) Without the appropriate state of mind there is no justification (so that external rightness is not a sufficient condition); and (2) a mistaken belief about the circumstances may justify (so that external rightness is not even a necessary condition).<sup>40</sup> Observe that these propositions would not justify the actions of the malicious, where their endeavors produced unintended but compensating benefits. Most of us, I suspect, would not con-

---

<sup>37</sup> FLETCHER, *RETHINKING CRIMINAL LAW*, *supra* note 1, at 767.

<sup>38</sup> Williams, *supra* note 1, *passim*; Robinson, *A Theory of Justification*, *supra* note 1, *passim*; Dressler, *Justifications and Excuses*, *supra* note 1, *passim*.

<sup>39</sup> See Greenawalt, *Perplexing Borders*, *supra* note 1; Greenawalt, *Distinguishing Justifications from Excuses*, *supra* note 1.

<sup>40</sup> Greenawalt, *Perplexing Borders*, *supra* note 1, at 1903-11.

sider that an unreasonable consequence. It means that we would have to punish the premeditated murder which unintentionally saved the lives of intended victims of the murdered person, and we would have to punish the unconsented-to borrowing of a car that resulted in saving the car from damage. It seems implausible, however, to allow the murderer or the thief to argue, "Look at the good I inadvertently did."

One of the cases at issue between the internalists and the externalists is the case of putative self-defense, in which the actor mistakenly believes she must harm another in self-defense. If the externalists, either pure or modified, are right, then putative self-defense is not a justification; but if the internalists are then putative self-defense is justified. One of the arguments raised by externalists is that one who injures another under the delusion of self-defense will feel remorse when she finds out the facts. But remorse is an indication, they say, that the action is at most to be excused, not justified. Remorse signifies a recognition that the deed was wrong. It is the role of excuse to relieve actors of liability for wrong actions that ought not to be attributed to them. Justified actions, on the other hand, are not wrong, and hence breed no remorse.<sup>41</sup>

Of this argument, Greenawalt says,

This argument underestimates the complexity of our moral practices. We learn to feel guilt over causing certain kinds of consequences, and we cannot turn off these feelings on the unusual occasions when our actions causing the consequences are warranted. . . . Feelings of remorse are quite compatible with claims that an action is justified.<sup>42</sup>

The argument is sound. If I am attacked by an old friend who is acting under the influence of a mental disease, I will be justified in resisting her; but I will certainly regret any harm I inflict on her. So a strong case may be made that the actor's perception of necessity is sufficient.

### 3. Conclusion: Rights of Third Parties

We have seen substantial support both for the claim that a justified action need only be a morally permissible or morally acceptable action, and for the claim that it is the actor's *perception* of permissibility or acceptability that matters. But these two claims undermine a third thesis about the distinction between justification and excuse, namely, that the fact of justification affects the rights of third parties in a way that excuse does not.

---

<sup>41</sup> See Fletcher, *Intolerable Prison Conditions*, *supra* note 1, at 1363.

<sup>42</sup> Greenawalt, *Perplexing Borders*, *supra* note 1, at 1921.

This thesis has been advanced by George Fletcher.<sup>43</sup> Fletcher reasoned that where Ploof is trying to tie up to Putnam's dock in order to avoid damage from a storm, neither Putnam nor anyone else is justified in stopping him if in fact Ploof is justified in tying up. But if Ploof is only excused for tying up, for example, on grounds of mental incompetence, then Putnam or a third party might well be justified in preventing him from doing so. This, he believes, reveals a fundamental difference between justification and excuse. We may be justified in resisting an excused action, but we can never be justified in resisting a justified action.

Fletcher's view is that a justified action is a right action, and it cannot also be right to prevent a right action. But if there are several justified actions and not just one (which there may be if the requirement for justification is that that action be morally permissible rather than morally required), why may a third party not resist one of those actions while helping to bring about another. For example, suppose Putnam, seeing Ploof about to dock, had seized the boat and tied it up at someone's else's slightly less valuable but equally sturdy dock. Although Ploof is not required to calculate the very best alternative among several that are reasonably similar, why should we prevent Putnam from doing it for him, if Putnam happens to know the relevant facts? That would seem to mean that one could be justified in resisting a justified action. Thus our conclusion that a justified action need not be unique contradicts this third thesis about the difference between justification and excuse.

The subjective characterization of justification is an even more devastating objection to Fletcher's rights-of-third-parties thesis. If what matters is not what alternatives actually face the actor, but what alternatives she *believes* face her, then it is possible for two people to be both justified while trying to bring about incompatible ends. For example, take a case of putative self-defense. If the actor mistakenly believes that she is under attack by someone, her defensive moves will be justified. At the same time, it will be justified for that other person to defend herself, and it might be justified for a third party to try to stop either of them.

Kent Greenawalt described the border between justification and excuse as "perplexed."<sup>44</sup> It is the commentators who are perplexed, and it seems to me that they are perplexed because they are looking for a border where none exists. Not that it is impossible to distinguish justification from other sorts of excuse. The difficulty

---

<sup>43</sup> FLETCHER, *RETHINKING CRIMINAL LAW*, *supra* note 1, at 760-61.

<sup>44</sup> Greenawalt, *Perplexing Borders*, *supra* note 1.

lies in trying to find a theoretical border between the two, a border that will show how the essence of justification is different from the essence of excuse. The two are essentially of the same sort. We justify and excuse for the same reason: because it would be unfair to do otherwise. The attempt to find a radically different basis for justification leads to Robinson's conclusion about the objectivity of justification—that people may be justified when they have the worst of intentions, and that they may not be justified when they are doing the best they can.<sup>45</sup> It is time, I think, to give up this quest, and to devote ourselves to the effort to limit the structure of excuses, including justifications.

#### B. THE SCHEME OF EXCUSES

All excuses rest in one way or another on the limitation of the actor's choices. The two general categories of excuse are deliberative exculpation and nondeliberative exculpation. Deliberative exculpation requires a choice on the part of the agent and a judgment by the trier of fact that the choice was acceptable. Deliberative exculpation is of two sorts, normative and causal. Nondeliberative exculpation requires that the actor be unable to make a reasonable choice. To claim nondeliberative exculpation is to claim that the appropriate course of action was simply not an alternative at all.

Deliberative exculpation includes both the standard deliberative excuses, like duress and necessity, and the justifications. A deliberative excuse, like duress, operates when the actor must choose under perceived pressure. For example, if A has threatened to kill B, a witness, unless B lies at A's trial, that provides B with no defense unless B is aware of it. The mere fact that B's options were restricted by A's threat is of no consequence if it was not the cause of B's behavior. B must be aware of the serious consequences that would attend obeying the law, and awareness of those consequences must influence her behavior.

How does her exculpation differ from the exculpation of one who is justified—one who trespasses, for example, to avoid harm to her own property? Is it that excuse depends on the internal state of the actor, while justification depends on some feature of the action? It is difficult to see how that could be; in each case, justification and excuse, it is the actor's perceptions that determine whether or not she has a defense. Nor, in the case of deliberative excuses, is there any reason to think that the actor is in any way deficient in her perception, or that it would matter if she were. We do not hold her to

---

<sup>45</sup> See Robinson, *A Theory of Justification*, *supra* note 1, *passim*.

an individual standard. She is measured against the person of "reasonable firmness."<sup>46</sup> Ordinary persons are excused if they act as ordinary persons would have acted, and they are justified if they act as ordinary persons would have. If you accept the subjective characterization of justification argued for above, then you will conclude that justification must be grouped with the deliberative excuses. Structurally the two are the same.

Thus, to claim deliberative exculpation is to claim that the actor was faced with a dilemma, and that the choice she made was an acceptable one. The distinction between normative and causal claims has to do not with the source of the restriction of the alternatives (with whether it was a natural occurrence or an intentional human action that created the dilemma), but rather with the nature of the appraisal. In normative exculpation we weigh the alternatives to determine which is the morally superior. In causal exculpation we are concerned with the difficulty of avoiding the different alternatives.<sup>47</sup>

A defendant is entitled to claim normative deliberative exculpation when, if matters were as she reasonably believed them to be, her action would be permissible. She may claim this even if conforming to the law were as easy as or easier than the action she has chosen. Her action will be, as Kent Greenawalt says, warranted.<sup>48</sup> She is entitled to claim causal deliberative exculpation when, if matters were as she reasonably believed them to be, conforming to the law would be so difficult as to make less difficult alternatives morally permissible and the alternative she has chosen is one of the permissible ones.

A defendant is entitled to claim nondeliberative exculpation

---

<sup>46</sup> MODEL PENAL CODE § 2.09.

<sup>47</sup> It is an interesting question whether anything can be made of duress *without* consideration of the difficulty of the alternatives. Suppose, for example, that a man is threatened with injury to his father if he does not participate in a bank holdup. As it happens, he has no love for his father and (being of rather low character) would not even mind seeing his father killed. In fact, he hopes to collect a large inheritance once his father dies. He is also not bothered by the idea that his refusal would bring about his father's injury and that in a way he would be responsible for it. Nevertheless, he agrees to the bank robbery (perhaps hoping to share in the loot). Obeying the law is not made more difficult for him here by the threat of injury to his father; in fact, obeying the law may present itself as the easier alternative. This man does not deserve to be excused. It is the difficulty of the alternatives that exculpates in situations of this sort, and only if the difficulty influences the action. This would be a case of duress and not of normative justification if the defense argued that the difficulty of seeing his father beaten caused him to take part in the robbery, and not that saving his father from harm was clearly superior to saving the bank. It makes a difference; in the second case the *foreseeable consequences* of robbing the bank would be relevant (would it somehow entail physical injury to many others?), but in the first they would not.

<sup>48</sup> Greenawalt, *Perplexing Borders*, *supra* note 1, at 1903.

when either the actor's perception or her ability to act on her perception is so impaired that it would be unreasonable to hold her to the usual standard of responsibility. The following scheme is meant to capture these distinctions:

- I. Deliberative Exculpatory Claims
  - A. Normative
    - 1. Justification
    - [2. Provocation]
  - B. Causal
    - 1. Duress
    - 2. The excuse of necessity
- II. Nondeliberative Exculpatory Claims
  - A. Disorders
    - 1. Insanity
    - 2. Involuntary Intoxication
    - 3. Infancy
  - [B. Ignorance and Mistake]

Provocation is not usually a defense, but rather a circumstance to be considered in mitigation. Still, the grounds upon which mitigation is based are plainly related to the grounds of exculpation. I have placed it with normative deliberative exculpation because that is where it belongs. The heart of provocation, as von Hirsch and Jareborg have argued, is not loss of control but rather *justified resentment*. Not everyone who becomes angry at another will have the mitigation of provocation if she breaks the law. In addition to acting in anger she must have "good reason for being angry in virtue of some wrong or impropriety suffered at [the provoker's] hands."<sup>49</sup> Rage alone, if they are right, will not amount to provocation, though at a certain level it might indeed provide an excuse of a different sort. While there are obviously other plausible views of this matter, I am persuaded by their argument, which I commend to the reader.

It is difficult to know where to place mistake and ignorance in this scheme. It may not be necessary to consider them separately at all. There are two ways in which mistake or ignorance might arise as a defense. The defendant might argue that she thought that she was doing a legally permitted act, as when someone shoots at what she takes to be a deer but is really a person. Second, she might argue that she thought that the alternatives to her illegal act would be calamitous. Self-defense is such a claim, as when she shoots a person thinking that that person is attacking her. To the extent that knowledge is an element of a crime, the first of these amounts to a denial of an element of a crime.<sup>50</sup> Mistake and ignorance in the second

<sup>49</sup> Hirsch & Jareborg, *supra* note 1, at 248.

<sup>50</sup> It does not follow inescapably from the fact that a claim is a denial of an element

sort of case become part of the deliberative exculpation in which they are asserted; in this case, self-defense (a variety of justification). In either case mistake does not require a special place in this scheme. Nevertheless, if mistake is to be included it would seem to fit best with the nondeliberative types of excuse, since it does not involve consideration of various alternatives, but rather a deficiency of a sort.

What are the practical consequences of the distinctions marked here? The consequences are those we are all already aware of. Justification depends on normative judgments about right and wrong, while duress depends on causal judgments. Another way to put it is that duress depends on normative judgments of a different sort, judgments about the reasonableness of requiring someone to undertake a certain level of difficulty.

One theory of excuses recently put forward finds the basis of excuse in interference with practical reasoning.<sup>51</sup> For one reason or another, the excused person is not able to reason effectively about her actions, or put her desires into effect. That basis of excuse works reasonably well if the paradigm for excuse is the nondeliberative excuse. But it does not fit duress for the same reason it does not fit justification. The person acting under duress need not lose control of her reasoning process; she might reason quite calmly in choosing among the available alternatives. In fact, to the extent she does not, her reasoning is confused because of the pressure she is under, and to that extent her excuse is not duress at all but something else entirely.

The correct theory about what underlies both justification and excuse will make exculpation depend on the limitation of alternatives open to the agent (a limitation caused by her circumstances), and on the permissibility of the alternative the agent has chosen.<sup>52</sup> Nondeliberative exculpation depends upon the agent being literally unable to make a different choice, a fact which by itself removes any moral or legal requirement that she act otherwise. Deliberative exculpation is different in that the agent is faced with a genuine choice. If exculpation is normative, then the alternative she has chosen is permissible because, in her understanding of things, the actual al-

---

that it should not be considered an excuse, and in fact some have asserted that it might be both denial of an element and an excuse. See Williams, *supra* note 1, at 735. Nevertheless that mistake should be one or the other but not both seems to me the better view of the matter, and I am aware of no unacceptable consequences of that view.

<sup>51</sup> MICHAEL MOORE, *LAW AND PSYCHIATRY*, *supra* note 1, at 84-90.

<sup>52</sup> In fact Michael Moore in his more recent writings has argued that his earlier "practical reasoning" strategy was simply a variety of the "she could not have done otherwise" defense. See Moore, *Choice, Character, and Excuse*, *supra* note 1, at 31.

ternative is genuinely preferable to its alternative (however easy or difficult that alternative may have been). If exculpation is causal, the alternative she has chosen is permissible because the difficulty of avoiding it makes it so.

Questions of public necessity which do not bear directly on the actor's fortunes are rated normatively. We compare what would be gained in obeying the law with what would be lost in obeying the law. The restriction to imminent harms insures that we use the appropriate mechanisms when possible without taking the law into our own hands. Nevertheless, I see no reason why this defense might not cover the emergency creation of benefits as well as the emergency avoidance of harm. There is always the tort remedy of compensation of course, and I see no reason why it is less just to allow us to use someone's property in an emergency to create a significant public benefit than to allow us to use it to prevent a significant public harm.

Questions of *personal* harm and benefit are different. Though we might excuse someone because a threat of personal harm may have amounted to duress, we do not excuse someone because a large bribe was just too large to pass up. Why should we feel that justification and excuse should lead to different conclusions in such a case?<sup>53</sup> The difference may lie in the distinction between normative and causal exculpation. While the normative has nothing to do with the difficulty of the alternatives (breaking the law may have been more difficult for the selfless hero than obeying it), difficulty is the very basis of causal exculpation. In the case of a bribe, the great difficulty is created by the size of the promised reward. At the same time, however, the attractiveness of that alternative (the difficulty of avoiding it) is seriously diminished by the expectation of punishment that accompanies the violation of the law. This shift in motivation does not have to be considered in the normative case. Threats, on the other hand, may exculpate in spite of the fear of the law. The determination the factfinder must make is precisely whether, in the reasonable person, fear of a particular bodily harm would rightly outweigh fear of punishment. When a threat does not exculpate, it is precisely because the fear of the threat would not outweigh the fear of the law. I think it is possible to imagine bribes so colossal that they would morally exculpate—though for policy reasons we might refuse to legally exculpate.

---

<sup>53</sup> You may, of course, avoid this difference in outcome by rejecting my controversial claim that the creation of positive benefits may work as a justification. Nevertheless, some account of the difference between threats and bribes is required by any theory of excuse that bases exculpation on the difficulty of the alternative facing the agent.